Training Large Language Models to Reason in a Continuous Latent Space

Abstract

Large language models (LLMs) are typically constrained to reason in the language space, where they express the reasoning process through a chain-of-thought (CoT) to solve complex problems. However, the language space may not always be optimal for reasoning. Most word tokens primarily ensure textual coherence and are not essential for reasoning, while some critical tokens require complex planning and pose challenges to LLMs. To explore the potential of reasoning beyond language, we introduce a new paradigm called Coconut (Chain of Continuous Thought). Coconut utilizes the last hidden state of the LLM as a representation of the reasoning state, termed"continuous thought."Instead of decoding this state into words, we feed it back to the model as the next input embedding directly in the continuous space. This latent reasoning paradigm enables an advanced reasoning pattern, where continuous thoughts can encode multiple alternative next steps, allowing the model to perform a breadth-first search (BFS) rather than committing prematurely to a single deterministic path as in CoT. Coconut outperforms CoT on logical reasoning tasks that require substantial search during planning and achieves a better trade-off between accuracy and efficiency.